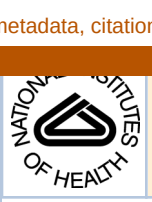

Published in final edited form as:
JAMA Psychiatry. 2013 March ; 70(3): 253–260. doi:10.1001/2013.jamapsychiatry.71.

A rare deletion at distal 16p11.2 is implicated in schizophrenia

Saurav Guha[1,§], **Elliott Rees**[2,§], **Ariel Darvasi**[3], **Dobril Ivanov**[2], **Masashi Ikeda**[4], **Sarah E. Bergen**[5,6,7], **Patrik K Magnusson**[8], **Paul Cormican**[9], **Derek Morris**[9], **Michael Gill**[9], **Sven Cichon**[10,11,12], **Jeffrey A. Rosenfeld**[1], **Annette Lee**[13], **Peter K. Gregersen**[13], **John M. Kane**[1], **Anil K. Malhotra**[1], **Marcella Rietschel**[14], **Markus M. Nöthen**[10,11,15], **Franziska Degenhardt**[10,11], **Lutz Priebe**[10,11], **René Breuer**[14], **Jana Strohmaier**[14], **Douglas M Ruderfer**[16], **Jennifer L Moran**[7], **Kimberly D Chambert**[7], **Alan R. Sanders**[17], **Jianxin Shi**[18], **Molecular Genetics of Schizophrenia Consortium**[19], **Wellcome Trust Case Control Consortium 2**[20], **Kenneth Kendler**[21], **Brien Riley**[22], **Tony O'Neill**[23], **Dermot Walsh**[24], **Dheeraj Malhotra**[25], **Aiden Corvin**[9], **Shaun Purcell**[5,6,7], **Pamela Sklar**[16], **Nakao Iwata**[4], **Christina M Hultman**[8], **Patrick F. Sullivan**[26], **Jonathan Sebat**[25], **Shane McCarthy**[27], **Pablo V. Gejman**[17], **Douglas F. Levinson**[28], **Michael J Owen**[2], **Michael C O'Donovan**[2], **Todd Lencz**[1,¶], and **George Kirov**[2,¶]

[1]The Zucker Hillside Hospital, Psychiatry Research, 75–59 263rd Street, Glen Oaks, NY 11004, USA [2]MRC Centre for Neuropsychiatric Genetics and Genomics, and Neuroscience and Mental Health Research Institute, Cardiff University, Heath Park, Cardiff CF4 4XN, UK [3]Department of Genetics, The Institute of Life Sciences, The Hebrew University of Jerusalem, Givat Ram, Jerusalem, Israel [4]Fujita Healt University School of Medicine, 1–98 Kutsukake-cho Dengakugakubo, Toyoake, Aichi, 470–1192, Japan [5]Psychiatric and Neurodevelopmental Genetics Unit, Department of Psychiatry, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA [6]Center for Human Genetics Research, Massachusetts General Hospital, Boston, MA, USA [7]Stanley Center for Psychiatric Research, Broad Institute, 7 Cambridge Center, Cambridge, MA, USA [8]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Sweden [9]Trinity College Dublin, Dept Psychiatry, College Green, Dublin, Ireland [10]Institute of Human Genetics, University of Bonn, Bonn, Germany [11]Department of Genomics, Life & Brain Center, University of Bonn [12]Institute of Neuroscience and Medicine (INM-1), Research Center Juelich, Juelich, Germany [13]Robert S. Boas Center for Human Genetics and Genomics, The Feinstein Institute for Medical Research, Manhasset, NY, USA [14]Dept. of Genetic Epidemiology in Psychiatry, Central Institute of Mental Health, Mannheim [15]German Center for Neurodegenerative Diseases (DZNE), Mount Siani school of Medicine, NY [16]Division of Psychiatric Genomics, Department Psychiatric, Mount Siani school of Medicine, NY [17]Department of Psychiatry and Behavioral Sciences, NorthShore University HealthSystem Research Institute, Evanston, IL & Department of Psychiatry and Behavioral Sciences, University of Chicago, Chicago, Illinois, USA [18]National Cancer Institute, Bethesda, MD [19]Molecular Genetics of Schizophrenia Consortium Affiliations are provided in the eSupplement, Section 9, Virginia Commonwealth University, Virginia, USA [20]Wellcome Trust Case Control Consortium 2 Affiliations are provided in the eSupplement, Section 10, Virginia Commonwealth University, Virginia, USA [21]Virginia Institute for Psychatry and Behavioral Genetics, School of Medicine,

[*]Corresponding authors: Todd Lencz, tlencz@NSHS.edu, and George Kirov, kirov@cardiff.ac.uk.
[§]these authors contributed equally to the research
[¶]these authors contributed equally to the research

Author Contributions

SG, ER, AD, MCO'D, TL, and GK wrote the manuscript. TL, AD, JMK, GK and AKM conceptualized and designed the study. SG, ER, TL, and GK performed the primary analyses, interpreted the genome wide data, and take responsibility for the integrity of the data and the accuracy of the data analysis. JR contributed in statistical analyses. All other authors contributed to collection, phenotyping, and genotyping of samples. All authors reviewed, edited, and approved the final manuscript.

Virginia Commonwealth University, Virginia, USA [22]Department of Psychiatry, School of Medicine, Virginia Commonwealth University, Virginia, USA [23]Department of Psychiatry, Queens University Belfast, Ireland [24]Health Research Board, Ireland [25]Beyster Center for Genomics of Psychiatric Diseases, University of California, San Diego, La Jolla, CA 92093, USA [26]Departments of Genetics, Psychiatry, and Epidemiology, University of North Carolina at Chapel Hill, USA [27]Stanley Institute for Cognitive Genomics. Woodbury Genome Center. Cold Spring Harbor Laboratory. Cold Spring Harbor, NY, USA [28]Department of Psychiatry and Behavioral Sciences, Stanford University, Stanford, CA

Abstract

Context—Large genomic copy number variations (CNVs) have been implicated as strong risk factors for schizophrenia. However, the rarity of these events has created challenges for the identification of further pathogenic loci, and extremely large samples are required to provide convincing replication.

Objective—To detect novel CNVs increasing susceptibility to schizophrenia, utilizing two ethnically homogeneous discovery cohorts and replication in large samples.

Design—Genetic association study of microarray data.

Setting—DNA samples were collected at nine sites from different countries.

Participants—Two discovery cohorts were comprised of: a) 790 cases (schizophrenia and schizoaffective disorder) and 1347 controls of Ashkenazi Jewish descent; and b) 662 trios (offspring affected with schizophrenia or schizoaffective disorder) from Bulgaria. Replication datasets consisted of 12,398 cases and 17,945 controls.

Main outcome measure—Statistically increased rate of specific CNVs in cases versus controls.

Results—One novel locus was implicated: a deletion at distal 16p11.2, which does not overlap the proximal 16p11.2 locus previously reported in schizophrenia and autism. Deletions at this locus were found in 13 out of 13,850 cases (0.094%) and in 3 out of 19,954 controls (0.015%), Fisher Exact p = 0.0014; OR = 6.25 (95%CI = 1.78 – 21.93).

Conclusion—Deletions at distal 16p11.2 have been previously implicated in developmental delay and obesity. The region contains nine genes, several of which are implicated in neurological diseases, regulation of body weight, and glucose homeostasis. A telomeric extension of the deletion, observed in about half the cases but no controls, potentially implicates an additional eight genes. Our findings add a new locus to the list of CNVs that increase risk to develop schizophrenia.

Introduction

Uncovering the genetic factors underlying schizophrenia (SZ) has proven difficult despite heritability estimates of up to 80% [1]. Copy number variants (CNVs) at several loci show consistently replicated evidence for association with SZ [2, 3]. These CNVs are individually very rare, are not fully penetrant, and are found cumulatively in ~2% of SZ cases; therefore, large samples were required to establish their association. Given their low baseline frequency, it is likely that further CNV susceptibility loci have yet to be discovered.

In the present study, we report the identification of a CNV locus at distal 16p11.2 that increases risk for SZ. Findings pointing to a possible association between this locus and SZ were obtained independently by two teams of investigators. During the process of obtaining

replication data, the two groups became aware of each other's work and decided to combine results from their discovery and replication cohorts. Using highresolution microarrays, one group (from New York and Israel) examined a SZ case-control cohort from the Ashkenazi Jewish (AJ) population, while the other group (from Cardiff, UK) examined a cohort of parent-offspring trios from Bulgaria (BG). Because of the need for large-scale replication, we contacted research groups worldwide willing to share raw data from microarray-based CNV studies in cohorts of SZ and control individuals, and obtained data from a total of ~34,000 individuals.

Methods

Bulgarian Trios sample (BG)

Sample description—The final sample (after QC) consisted of 662 Bulgarian offspring with all their parents, in 638 families (615 families with one offspring, 22 with two offspring and one with three). Details on this cohort have been previously described[4], but that previous publication only reported on *de novo* CNVs; here we report on the transmitted CNVs in this cohort. This cohort does not include patients with severe developmental disorders (all probands had attended mainstream schools, from which people with known intellectual disability were excluded). Diagnoses were made according to DSM-IV criteria [5], using a SCAN [6] interview and review of hospital discharge summaries. We included patients with schizophrenia or schizoaffective disorder. Concomitant medical conditions were not systematically assessed, except as related to psychiatric diagnosis. The CNVs found in the parents of each trio but not transmitted to the affected offspring comprised the "pseudocontrol" population listed in Table 1 under "controls".

Genotyping and Quality Control (QC)—All samples were genotyped on Affymetrix 6.0 arrays at the Broad Institute, USA. Analysis was performed using Genotyping Console 4.0 software, one batch at a time, with each batch containing 70–90 arrays. QC included removal of CNVs if they were from the X or Y chromosome, less than 15kb, covered by less than 15 probes or a probe density (size/probe number) greater than 7500bp. PLINK v1.07 [7] was used to exclude CNVs if 50% or more of their length was covered by a segmental duplication (SD). CNV loci with a frequency greater than 1% were excluded. Individuals with multiple large duplicate CNVs on the same chromosome were excluded, as these are likely to be artifacts [8]. Samples were also removed if their total number of CNVs was very high and constituted an outlier for the distribution within that sample (>50 CNVs for this experiment).

For additional QC of the Bulgarian trios we used a modification of the MeZOD algorithm proposed by McCarthy et al 2009 [9] and described in detail in Kirov et al., 2012 [4]. A Z-Score is the median of the standardised Log2 ratios for all probes within a specified chromosomal region. Through comparison of all individual Z-Scores for a given region, true CNVs are represented as outliers from the Z-Scores normal distribution. We show the distribution of the z-scores for the 16p11.2 distal region in the eSupplement (eFigure 1), which demonstrate that the only outliers for this region are the two probands with deletions, and their parents.

Ashkenazi Jewish sample (AJ)

Sample description—Case (n=1156) and control (n=2279) samples were selected from an Ashkenazi Jewish repository (Hebrew University Genetic Resource, HUGR, http://hugr.huji.ac.il). Patients for discovery analysis were recruited from hospitalized inpatients at seven medical centres in Israel. All diagnoses were assigned after direct interview using the structured clinical interview (SCID) [10], a questionnaire with inclusion and exclusion criteria, and cross-references to medical records. Chronic medical disorders and conditions were

recorded based on both patient report and hospital records. The inclusion criteria specified that subjects had to be diagnosed with SZ or schizoaffective disorder by DSM-IV criteria [5], that all four grandparents of each subject were reported to be of Ashkenazi Jewish ethnic origin, and that each subject or the subject's legal representative has signed the informed-consent form. Exclusion criteria included psychotic disorder due to a general medical condition, substance-induced psychotic disorder, pervasive developmental disorders, or any Cluster A (schizotypal, schizoid, or paranoid) personality disorder. Samples from healthy Ashkenazi individuals were collected from volunteers at the Israeli Blood Bank; these subjects were not psychiatrically screened but reported no chronic disease and were taking no medication at the time of blood draw. Corresponding institutional review boards and the National Genetic Committee of the Israeli Ministry of Health approved the studies. All samples were fully anonymized immediately after collection and subsequently, genomic DNA was extracted from blood samples through use of the Nucleon kit (Pharmacia). Genotyping and analyses were performed under protocols approved by the Institutional Review Board of the North Shore-LIJ Health System.

Genotyping and Quality Control (QC)—Genotyping was performed with Illumina HumanOmni1-Quad arrays according to manufacturers' specifications for ~ 1.4 million genome wide markers (~900K SNPs and ~500K CNV intensity probes). SNPs were filtered on the following basis: call rate < 98%, minor allele frequency < 0.02 and Hardy-Weinberg exact test $P < 0.000001$ in controls. Samples were filtered based on genotype quality control filtration (sample call rate < 97 %, gender mismatch) and examined for cryptic identity and first- or second-degree relatedness using pairwise identity-by-decent (IBD) estimation (PI_HAT) in PLINK [7] with 128,403 LD pruned ($r^2 > 0.2$) genome wide SNPs. Samples were excluded based on PI_Hat > 0.125; the individual with the lower call rate from each control/control or case/case pair was excluded, and controls were excluded from case/control pairs. The remaining samples were further examined for underlying population stratification using Principal Component Analysis (PCA) with Ancestry Informative Markers (AIMs) specific for the Ashkenazi Jewish population [11]. Samples with PCA results suggestive of one or more non-AJ grandparents were identified as outliers based on first principal component score > 0.01 and were excluded from further analysis (eSupplement, eFigure2). After quality control based on SNP markers, the dataset contained 2544 samples comprised of 904 cases (573 male and 331 female) and 1640 controls (1216 male and 424 female) genotyped on 762,372 high-quality SNPs with 99.8 % overall call rate.

CNV calls and validation—Normalization and log ratio data calculation for 904 cases and 1640 controls were performed using Illumina GenomeStudio. The resulting log2 R ratios (LRR) and B-allele frequencies (BAF) were used to identify CNVs on autosomes for each subject. We used variations of three algorithms for CNV detection: PennCNV [8], QuantiSNP [12] and cnvPartition (www.illumina.com). QuantiSNP and PennCNV are based on Hidden markov model (HMM) and cnvPartition is based on bivariate Gaussian distribution as implemented at Illumina GenomeStudio (www.illumina.com).

Following the methods of Need et al., (2009) [13] and Sanders et al. (2011) [14], we excluded any individuals with pennCNV threshold of LogR standard deviation (LRR_SD) ≥ 0.30, BAF drift ≥ 0.002, and/or Waviness factor (WF) deviating from 0 by > 0.04. Individuals containing > 500 CNVs (before filtration described below) were also excluded from the analysis. The final dataset contains 790 cases and 1347 controls.

We further excluded CNV calls based on QC thresholds recommended by each of the respective algorithms. Thus, CNV calls were excluded from further analysis if the Log Bayes Factor was ≤ 10 in QuantiSNP, confidence threshold ≤ 35 in cnvPartition, or default QC parameters in PennCNV were not obtained.

Following the QC steps all the CNV calls were merged using CNVision program [14] The final rare CNV calls were made based on consensus calls from all three algorithms (with no more than 25% of the length drawn from one algorithm only), with the following filtration criterion: ≥ 20 probes, ≥ 100kb in size, and <1% frequency in the total sample. CNVs of the same type (i.e., deletion or duplication) that were separated by ≤3 probes were merged into one contiguous segment as recommended by Vacic et al., 2011 [15]. All CNVs were annotated using CNVision. Based on previous findings in SZ and other neuropsychiatric disorders[16], purely intergenic CNVs were excluded.

Replication samples

Evidence for replication of the findings was obtained from seven case-control samples recruited and genotyped by other teams from the USA, Europe and Japan. These comprised 12,398 cases and 17,945 controls genotyped with high-resolution arrays. Details on the samples, genotyping platforms and QC used by these teams are detailed in the eSupplement, Section 1. The minimally affected region was covered well by all arrays used by the other teams (eFigure 3).

All coordinates in this paper are based on the human genome build NCBI36/hg18.

Results

After stringent quality control procedures, 790 cases and 1347 controls from the AJ cohort, and 662 probands from 638 BG families were examined for rare, large CNVs. Replication was sought in other case-control datasets for any CNVs that were observed in at least two cases and no controls (for the AJ cohort) and transmitted at least twice with no non-transmissions in the BG trios cohort. Several relevant CNVs were found in the two discovery datasets at loci already reported to increase risk to develop SZ, but as they are already known susceptibility factors, we only list them in the eSupplement eTable 4. In the AJ cohort, CNVs at two additional loci were observed in two cases and no controls. These were at chromosome 6q14.3 (hg18 coordinates: 85.25–85.58Mb) and 7q33 (133.39–133.50Mb), but replication evidence was not observed. No other CNV of this type was supported by replication evidence in the BG data (apart from the 16p11.2 deletion). The lists of all rare and large (>100kb) CNVs, in the two samples, that intersected genes, are available as eSupplement files (AJ_SZ_CNVs_over_100kb.xls and BG_SZ_trios_CNVs_over_100kb.xls).

The only CNV of interest that overlapped between the two discovery samples was a deletion at the distal region of 16p11.2, with a minimal common region between 28.73–28.95Mb (build 36, hg18). This region intersects nine genes and is flanked by two SD blocks (Figure 1). It does not overlap the known 16p11.2 locus at 29.56–30.11Mb that has been implicated in SZ [9, 16], autism [14, 17] and developmental delay [18]. Deletions at this locus were found in two cases (and no controls) from the AJ cohort and two offspring from the BG samples, both transmitted from mothers (there were no parents who did not transmit this CNV). Duplications at this locus were observed in one AJ control and in one BG parent (who transmitted it to an affected offspring).

We sought evidence for association between this deletion with SZ in seven independent case-control cohorts (12,398 cases and 17,945 controls) where we had access to the raw data (Table 1 and eSupplement, Section 1). Deletions overlapping this region were observed in an additional nine cases and three controls (Fisher exact for the replication sample p = 0.018, one-tailed; OR = 4.35 (95% CI = 1.18 - 16.06). Combining the discovery and replication cohorts, we found 13 deletions among 13,850 cases (0.094%) and three among 19,954 controls (0.015%) (two-tailed Fisher exact test p = 0.0014, OR = 6.25, 95%CI = 1.78 -

21.93). The positions of the CNVs are shown in Figure 1. There was no excess of duplications in cases at distal 16p11.2.

The minimal common region for all deletions reported in Table 1 encompasses nine genes within a 220kb interval flanked by blocks of SD (Figure 1). Some CNVs extend over the SDs, (however, we note that no CNV in the 16p11.2 region was excluded on the basis of >50% overlap with SD). Different breakpoints that extend over the flanking SD regions (but do not reach the telomeric region that is free of SDs) are more likely to reflect the different coverage of arrays (eSupplement eFigure 3) and/or the problems of calling CNVs over repetitive regions, rather than to have different pathogenicity, especially as these regions have fewer genes. Seven deletions cover an additional region of unique DNA sequence, at the telomeric side (very left on Figure 1, the interval free of SDs), that contains further genes. Evidence for pathogenicity of the seven CNVs that extended over the telomeric region was nearly as strong as for the implicated critical region (7/13,850 cases vs. 0/19,954 controls, two-tailed Fisher exact test p = 0.0019). However the critical "distal 16p11.2 region" remains the more likely candidate due to its confirmed involvement in other neurodevelopmental disorders (see Discussion), and the lack of isolated CNVs in the smaller telomeric region. Out of the three controls with deletions, one (in the Swedish dataset) was recruited at the age of 45, had diabetes type 2, and high blood pressure, but no other medical or psychiatric problems. No further information is available on the two anonymized controls from the WTCCC2/Irish dataset: one is from the British Blood Transfusion service (therefore presumably healthy), and the other one from the 1958 cohort.

Importantly, the new "distal" locus is approximately 600kb telomeric from the previously implicated "proximal" 16p11.2 CNV (29.56–30.11Mb) [9]. CNVs at "proximal" 16p11.2 have been shown to increase risk for SZ, autism, and developmental delay when duplicated [5,16], and for autism and developmental delay when deleted [9, 14, 18]. None of the CNVs in our study extend over the "proximal" region (Figure 1).

We have previously demonstrated that the known SZ-associated CNVs have high mutation rates and that strong selection pressure operates against them [24]. We are able to estimate the *de novo* rate for this deletion at 25% based on the current study (two transmitted deletions and no information on inheritance in the other subjects) and four available datasets with a total of 5 de novo occurrences out of 20 events with a known inheritance [19–22] (eSupplement Section 7). This approximates to a selection pressure of 0.25. In line with this we observe the two BG proband deletions to be found on different haplotypes, and therefore very likely to be independent mutations.

Phenotypic data, where available, indicate a spectrum of typical presentations of SZ with no evidence for intellectual disability, or a specific clinical profile (eTable 2). This is similar to the lack of specific clinical presentations reported for the other large CNVs implicated in neurodevelopmental disorders [2, 14, 19]. The possible exception is the presence of two individuals with obesity and two with type 2 diabetes (plus one control with type 2 diabetes) in line with previous reports, (see Discussion). In addition to the 13 cases listed in Table 1 and eTable 2, we note that the brother of one case (in the Japanese sample) carries the same deletion and is also affected with SZ. Further probands had positive family histories of SZ, but we do not know if their affected relatives also carry the deletion. Although the transmission status of the CNVs is only available for the BG cohort, we further note that both deletions were transmitted maternally.

Discussion

Several lines of evidence from the literature support the distal 16p11.2 deletion as a true SZ-associated CNV locus. The deletion has been implicated in developmental delay and other clinical phenotypes [18, 19, 20] (details in eTable 1), similar to other SZ-associated CNVs [2, 3, 9]. Briefly, Cooper et al. (2011) [18] reported a very similar increased rate of 0.1% (15/15,767) for this deletion in children with intellectual disability, autism spectrum disorders and congenital malformations, that were referred for genetic testing, compared with a control rate of 0.01% (1/8329). Similar rates were found in another large study on patients with developmental delay and a range of other abnormal phenotypes [19]: 31/23,084 cases (0.13%) and 1/7700 controls (0.01%). Interestingly, out of the six cases in that study, for whom detailed clinical information was available, one had autism, behavioral problems/ADHD and SZ, another one had behavioral problems/ADHD and bipolar disorder, and a third one had autism. Four of these six cases were overweight and all six had developmental delay. Moreover, additional telomeric extension of the deletion (to approx. 28.4Mb) was present in 9 of the 31 cases and was never observed in controls. Similarly, in our study, 7 of the 13 of the cases demonstrated this telomeric extension, whereas this was not seen in the controls. We note that the controls used in these studies partially overlap ours, so these control rates are not independent, (eSupplement Section 4, eTable 1). Additional published reports of distal 16p11.2 deletions include five patients from two separate families [21], all of whom have developmental delay and behavioral problems, and one child out of 4284 patients with mental retardation [22].

Distal 16p11.2 deletions have also been shown to be enriched in patients with severe early-onset obesity (3/300 = 1%) compared to unscreened population controls (2/7366 = 0.03%) [20], consistent with the findings in the study by Bachmann-Gagescu et al [19], discussed above. It was postulated that the most likely obesity candidate within the distal 16p11.2 region is *SH2B1*, as this gene plays a role in the regulation of body weight and glucose homeostasis in mice [23]. Two of our cases were obese/overweight, and two cases and one control had type 2 diabetes, (consistent with being overweight although this information is not available). However, one carrier (from Japan) had documented evidence of normal weight, and several did not have recorded evidence of obesity despite being drawn from cohorts that were assessed for this and other medically-relevant phenotypes.

Considerable heterogeneity of phenotypic expression has been reported for most large rare CNVs implicated in SZ, with carriers often manifesting non-psychotic phenotypes including intellectual disability, autism, epilepsy, obesity, and cardiac disorders [2,16]. Pleiotropy appears to also be the case for distal 16p11.2 deletions, possibly due to the presence of multiple genes within the deleted region.

Clinical presentations for distal 16p11.2 deletion carriers are unremarkable for SZ, with diagnoses ranging across all major subtypes: paranoid, catatonic, undifferentiated and schizoaffective. Age of SZ onset for deletion carriers ranges from 15–30 (mean = 23.4), with no clear evidence for early onset. Two parents who transmitted the deletions to probands did not have psychotic disorders, although one had a mood disorder. Out of the three controls who carry the deletion, one (from Sweden) did not report psychiatric problems at the age of 45, when interviewed, past the usual accepted age for the period of risk for SZ. Of the other two controls, one had also passed through the risk period (from the 1958 cohort, examined at the age of 44–45, see eSupplement) and the third one, a blood donor, is presumably healthy and not on any medication. These observations indicate that this CNV does not have full penetrance, similar to most other CNVs implicated in SZ. None of the carriers had any other SZ-associated CNVs.

Mutations in several of the nine genes within the critical region of distal 16p11.2 have been implicated in neurological diseases: homozygous mutations in the gene *TUFM* have been described in infants with fatal encephalopathy [25]; *ATP2A1* is implicated in Brody disease in which patients are unable to relax their muscle during exercise [26], and its homologue, *ATP2A2* has been implicated in neuropsychiatric phenotypes [27]; *ATXN2L* (although unknown in function) encodes a protein belonging to the spinocerebellar ataxia family. The remaining genes are either involved in immunity, insulin and leptin signaling (*SH2B1*) or are of unknown function. In addition to the nine genes in the minimal critical region, the larger CNVs with telomeric extensions include eight additional deleted genes (seven of them in DNA region that is free of SDs, Figure 1), possibly increasing the pathogenicity of these larger CNV. Most notable among these eight genes is *CLN3*, where recessive mutations are associated with Batten disease, characterized by childhood-onset neurodegeneration [28]. Moreover, *CLN3* is the only gene in either the minimal or the extended region that is implicated in synaptic function based on Gene Ontology annotation. Our previous study of *de novo* CNVs indicated an enrichment of such genes in SZ-related events, however *CLN3* is not among the post-synaptic density (PSD) genes, implicated in that study [4]. Additional evidence from animal knockout models may help disentangle the contributions of each of these genes to the observed range of phenotypes.

In conclusion, we have obtained strong evidence for the role of a new CNV locus in SZ. Similar to other such loci, it is very rare and increases risk for other neurodevelopmental phenotypes.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Work at Cardiff University was funded by Medical Research Council (MRC) Programme (Ref G0800509) and Centre Grants, the National Institutes of Mental Health (USA) (CONTE: 2 P50 MH066392–05A1), and a grant from the EU (EU-GEI). Genotyping was funded by multiple grants to the Stanley Center for Psychiatric Research at the Broad Institute from the Stanley Medical Research Institute, The Merck Genome Research Foundation, and the Herman Foundation. The Zucker Hillside work was supported by the North Shore-LIJ Health System Foundation and the National Institutes of Health (RC2 MH089964 and R01 MH084098 to T.L; P50 MH080173 to A.K.M and P30 MH090590 to J.M.K).

MGS data collection was supported by National Institute of Mental Health (NIMH) R01 grants MH67257, MH59588, MH59571., MH59565, MH59587, MH60870, MH59566, MH59586, MH61675, MH60879, and MH81800, NIMH U01 grants MH46276, MH46289, MH46318, MH79469 and MH79470. Support is acknowledged from NARSAD (National Alliance for Research on Schizophrenia and Depression) Young Investigator Awards (to J.D. and A.R.S.), the Genetic Association Information Network (GAIN), the Paul Michael Donovan Charitable Foundation, and the Walter E. Nichols, M.D., and Eleanor Nichols endowments (Stanford University). Genotyping was carried out by the Center for Genotyping and Analysis at the Broad Institute of Harvard and MIT (S. Gabriel and D. B. Mirel) with support from grant U54 RR020278 from the National Center for Research Resources.

The German study was supported by the German Federal Ministry of Education and Research (BMBF), within the context of the Integrated Genome Research Network (IG) MooDS (grants 01GS08144 and 01GS08147). Control data were provided by the community-based studies of PopGen, KORA, the Heinz Nixdorf Recall (HNR) study, and MooDS.

Genotyping of WTCCC2 samples was funded by the Wellcome Trust (085475/B/08/Z and 085475/Z/08/Z). The British 1958 Birth Cohort DNA collection was funded by the UK Medical Research Council (G0000934) and the Wellcome Trust (068545/Z/02), and the UK National Blood Service controls by the Wellcome Trust.

The Japanese cohort was supported in part by research grants from the Japan Ministry of Education, Culture, Sports, Science and Technology; the Ministry of Health, Labor and Welfare; the Core Research for Evolutional

Science and Technology; and the Health Sciences Foundation (Research on Health Sciences focusing on Drug Innovation).

Funding sources did not have any role in the design of study and approval of manuscript.

The authors report no relevant financial interests to disclose.

References

1. Owen MJ, Williams NM, O'Donovan MC. The molecular genetics of schizophrenia: new findings promise new insights. Mol Psychiatry. 2003; 9:14–27. [PubMed: 14581932]
2. Kirov G. The role of copy number variation in schizophrenia. Expert Review of Neurotherapeutics. 2010; 10:25–32. [PubMed: 20021318]
3. Levinson DF, Duan J, Oh S, Wang K, Sanders AR, Shi J, Zhang N, Mowry BJ, Olincy A, Amin F, Cloninger CR, Silverman JM, Buccola NG, Byerley WF, Black DW, Kendler KS, Freedman R, Dudbridge F, Pe'er I, Hakonarson H, Bergen SE, Fanous AH, Holmans PA, Gejman PV. Copy number variants in schizophrenia: Confirmation of five previous findings and new evidence for 3q29 microdeletions and VIPR2 duplications. Am J Psychiat. 2011; 168:302–316. [PubMed: 21285140]
4. Kirov G, Pocklington AJ, Holmans P, Ivanov D, Ikeda M, Ruderfer D, Moran J, Chambert K, Toncheva D, Georgieva L, Grozeva D, Fjodorova M, Wollerton R, Rees E, Nikolov I, Lagemaat LN, Bayés A, Fernandez E, Olason PI, Böttcher Y, Komiyama NH, Collins MO, Choudhary J, Stefansson K, Stefansson H, Grant SG, Purcell S, Sklar P, O'Donovan MC, Owen MJ. De novo CNV analysis implicates specific abnormalities of postsynaptic signalling complexes in the pathogenesis of schizophrenia. Mol Psychiatr. 2012; 17:142–153.
5. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. Fourth. American Psychiatric Association; Washington, DC: 1994.
6. Wing JK, Babor T, Brugha T, Burke J, Cooper JE, Giel R, Jablenski A, Regier D, Sartorius N. SCAN: Schedules for Clinical Assessment in Neuropsychiatry. Arch Gen Psychiatry. 1990; 47:589–593. [PubMed: 2190539]
7. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC. PLINK: A tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007; 81:559–575. [PubMed: 17701901]
8. Wang K, Li M, Hadley D, Liu R, Glessner J, Grant SF, Hakonarson H, Bucan M. PennCNV: An integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data. Genome Res. 2007; 17:1665–1674. [PubMed: 17921354]
9. McCarthy SE, Makarov V, Kirov G, Addington AM, McClellan J, Yoon S, Perkins DO, Dickel DE, Kusenda M, Krastoshevsky O, Krause V, Kumar RA, Grozeva D, Malhotra D, Walsh T, Zackai EH, Kaplan P, Ganesh J, Krantz ID, Spinner NB, Roccanova P, Bhandari A, Pavon K, Lakshmi B, Leotta A, Kendall J, Lee YH, Vacic V, Gary S, Iakoucheva LM, Crow TJ, Christian SL, Lieberman JA, Stroup TS, Lehtimäki T, Puura K, Haldeman-Englert C, Pearl J, Goodell M, Willour VL, Derosse P, Steele J, Kassem L, Wolff J, Chitkara N, McMahon FJ, Malhotra AK, Potash JB, Schulze TG, Nöthen MM, Cichon S, Rietschel M, Leibenluft E, Kustanovich V, Lajonchere CM, Sutcliffe JS, Skuse D, Gill M, Gallagher L, Mendell NR, Wellcome Trust Case Control Consortium. Craddock N, Owen MJ, O'Donovan MC, Shaikh TH, Susser E, Delisi LE, Sullivan PF, Deutsch CK, Rapoport J, Levy DL, King MC, Sebat J. Microduplications of 16p11.2 are associated with schizophrenia. Nat Genet. 2009; 41:1223–1227. [PubMed: 19855392]
10. Spitzer RL, Williams JBW, Gibbon M, First MB. The Structured Clinical Interview for DSM-III-R (SCID) I: History, rationale, and description. Arch Gen Psychiatry. 1992; 49:624–629. [PubMed: 1637252]
11. Guha S, Rosenfeld JA, Malhotra AK, Lee AT, Gregersen PK, Kane JM, Pe'er I, Darvasi A, Lencz T. Implications for health and disease in the genetic signature of the Ashkenazi Jewish population. Genome Biol. 2012; 13(1):R2. Epub ahead of print. [PubMed: 22277159]
12. Colella S, Yau C, Taylor JM, Mirza G, Butler H, Clouston P, Bassett AS, Seller A, Holmes CC, Ragoussis J. QuantiSNP: an Objective Bayes Hidden-Markov Model to detect and accurately map

copy number variation using SNP genotyping data. Nucleic Acids Res. 2007; 35:2013–2025. [PubMed: 17341461]

13. Need AC, Ge D, Weale ME, Maia J, Feng S, Heinzen EL, Shianna KV, Yoon W, Kasperaviciūte D, Gennarelli M, Strittmatter WJ, Bonvicini C, Rossi G, Jayathilake K, Cola PA, McEvoy JP, Keefe RS, Fisher EM, St Jean PL, Giegling I, Hartmann AM, Möller HJ, Ruppert A, Fraser G, Crombie C, Middleton LT, St Clair D, Roses AD, Muglia P, Francks C, Rujescu D, Meltzer HY, Goldstein DB. A genome-wide investigation of SNPs and CNVs in schizophrenia. PLoS Genet. 2009; 5:e1000373. [PubMed: 19197363]
14. Sanders SJ, Ercan-Sencicek AG, Hus V, Luo R, Murtha MT, Moreno-De-Luca D, Chu SH, Moreau MP, Gupta AR, Thomson SA, Mason CE, Bilguvar K, Celestino-Soper PB, Choi M, Crawford EL, Davis L, Wright NR, Dhodapkar RM, DiCola M, DiLullo NM, Fernandez TV, Fielding-Singh V, Fishman DO, Frahm S, Garagaloyan R, Goh GS, Kammela S, Klei L, Lowe JK, Lund SC, McGrew AD, Meyer KA, Moffat WJ, Murdoch JD, O'Roak BJ, Ober GT, Pottenger RS, Raubeson MJ, Song Y, Wang Q, Yaspan BL, Yu TW, Yurkiewicz IR, Beaudet AL, Cantor RM, Curland M, Grice DE, Günel M, Lifton RP, Mane SM, Martin DM, Shaw CA, Sheldon M, Tischfield JA, Walsh CA, Morrow EM, Ledbetter DH, Fombonne E, Lord C, Martin CL, Brooks AI, Sutcliffe JS, Cook EH Jr, Geschwind D, Roeder K, Devlin B, State MW. Multiple recurrent de novo CNVs, including duplications of the 7q11.23 Williams Syndrome region, are strongly associated with autism. Neuron. 2011; 70:863–885. [PubMed: 21658581]
15. Vacic V, McCarthy S, Malhotra D, Murray F, Chou HH, Peoples A, Makarov V, Yoon S, Bhandari A, Corominas R, Iakoucheva LM, Krastoshevsky O, Krause V, Larach-Walters V, Welsh DK, Craig D, Kelsoe JR, Gershon ES, Leal SM, Dell Aquila M, Morris DW, Gill M, Corvin A, Insel PA, McClellan J, King MC, Karayiorgou M, Levy DL, DeLisi LE, Sebat J. Duplications of the neuropeptide receptor gene VIPR2 confer significant risk for schizophrenia. Nature. 2011; 471:499–503. [PubMed: 21346763]
16. Malhotra D, Sebat J. CNVs: Harbingers of a Rare Variant Revolution in Psychiatric Genetics. Cell. 2012; 148:1223–41. [PubMed: 22424231]
17. Levy D, Ronemus M, Yamrom B, Lee YH, Leotta A, Kendall J, Marks S, Lakshmi B, Pai D, Ye K, Buja A, Krieger A, Yoon S, Troge J, Rodgers L, Iossifov I, Wigler M. Rare de novo and transmitted copy-number variation in autistic spectrum disorders. Neuron. 2011; 70:886–97. [PubMed: 21658582]
18. Cooper GM, Coe BP, Girirajan S, Rosenfeld JA, Vu TH, Baker C, Williams C, Stalker H, Hamid R, Hannig V, Abdel-Hamid H, Bader P, McCracken E, Niyazov D, Leppig K, Thiese H, Hummel M, Alexander N, Gorski J, Kussmann J, Shashi V, Johnson K, Rehder C, Ballif BC, Shaffer LG, Eichler EE. A copy number variation morbidity map of developmental delay. Nat Genet. 2011; 43:838–846. [PubMed: 21841781]
19. Bachmann-Gagescu R, Mefford HC, Cowan C, Glew GM, Hing AV, Wallace S, Bader PI, Hamati A, Reitnauer PJ, Smith R, Stockton DW, Muhle H, Helbig I, Eichler EE, Ballif BC, Rosenfeld J, Tsuchiya KD. Recurrent 200-kb deletions of 16p11.2 that include the SH2B1 gene are associated with developmental delay and obesity. Genet Med. 2010; 12:641–647. [PubMed: 20808231]
20. Bochukova EG, Huang N, Keogh J, Henning E, Purmann C, Blaszczyk K, Saeed S, Hamilton-Shield J, Clayton-Smith J, O'Rahilly S, Hurles ME, Farooqi IS. Large, rare chromosomal deletions associated with severe early-onset obesity. Nature. 2010; 463:666–670. [PubMed: 19966786]
21. Barge-Schaapveld DQ, Maas SM, Polstra A, Knegt LC, Hennekam RC. The atypical 16p11.2 deletion: A not so atypical microdeletion syndrome? Am J Med Genet A. 2011; 155:1066–1072. [PubMed: 21465664]
22. Bijlsma EK, Gijsbers AC, Schuurs-Hoeijmakers JH, van Haeringen A, Fransen van de Putte DE, Anderlid BM, Lundin J, Lapunzina P, Pérez Jurado, LA Delle, Chiaie B, Loeys B, Menten B, Oostra A, Verhelst H, Amor DJ, Bruno DL, van Essen AJ, Hordijk R, Sikkema-Raddatz B, Verbruggen KT, Jongmans MC, Pfundt R, Reeser HM, Breuning MH, Ruivenkamp CA. Extending the phenotype of recurrent rearrangements of 16p11.2: Deletions in mentally retarded patients without autism and in normal individuals. Eur J Med Genet. 2009; 52:77–87. [PubMed: 19306953]

23. Morris DL, Cho KW, Rui L. Critical role of the Src Homology 2 (SH2) domain of neuronal SH2B1 in the regulation of body weight and glucose homeostasis in mice. Endocrinology. 2010; 151:3643–3651. [PubMed: 20484460]
24. Rees E, Moskvina V, Owen MJ, O'Donovan MC, Kirov G. De novo rates and selection of schizophrenia-associated copy number variants. Biol Psychiat. 2011; 70:1109–1114. [PubMed: 21855053]
25. Valente L, Tiranti V, Marsano RM, Malfatti E, Fernandez-Vizarra E, Donnini C, Mereghetti P, De Gioia L, Burlina A, Castellan C, Comi GP, Savasta S, Ferrero I, Zeviani M. Infantile encephalopathy and defective mitochondrial DNA translation in patients with mutations of mitochondrial elongation factors EFG1 and EFTu. Am J Hum Genet. 2007; 80:44–58. [PubMed: 17160893]
26. Odermatt A, Taschner PE, Khanna VK, Busch HF, Karpati G, Jablecki CK, Breuning MH, MacLennan DH. Mutations in the gene-encoding SERCA1, the fast-twitch skeletal muscle sarcoplasmic reticulum Ca2+ ATPase, are associated with Brody disease. Nat Genet. 1996; 14:191–194. [PubMed: 8841193]
27. Gordon-Smith K, Jones LA, Burge SM, Munro CS, Tavadia S, Craddock N. The neuropsychiatric phenotype in Darier disease. Brit J Dermatol. 2010; 163:515–522. [PubMed: 20456342]
28. Lerner TJ, Boustany RN, Anderson JW, D'Arigo KL, Schlumpf K, Buckler AJ, Gusella JF, Haines JL. Isolation of a novel gene underlying batten disease, CLN3. Cell. 1995; 82:949–957. [PubMed: 7553855]

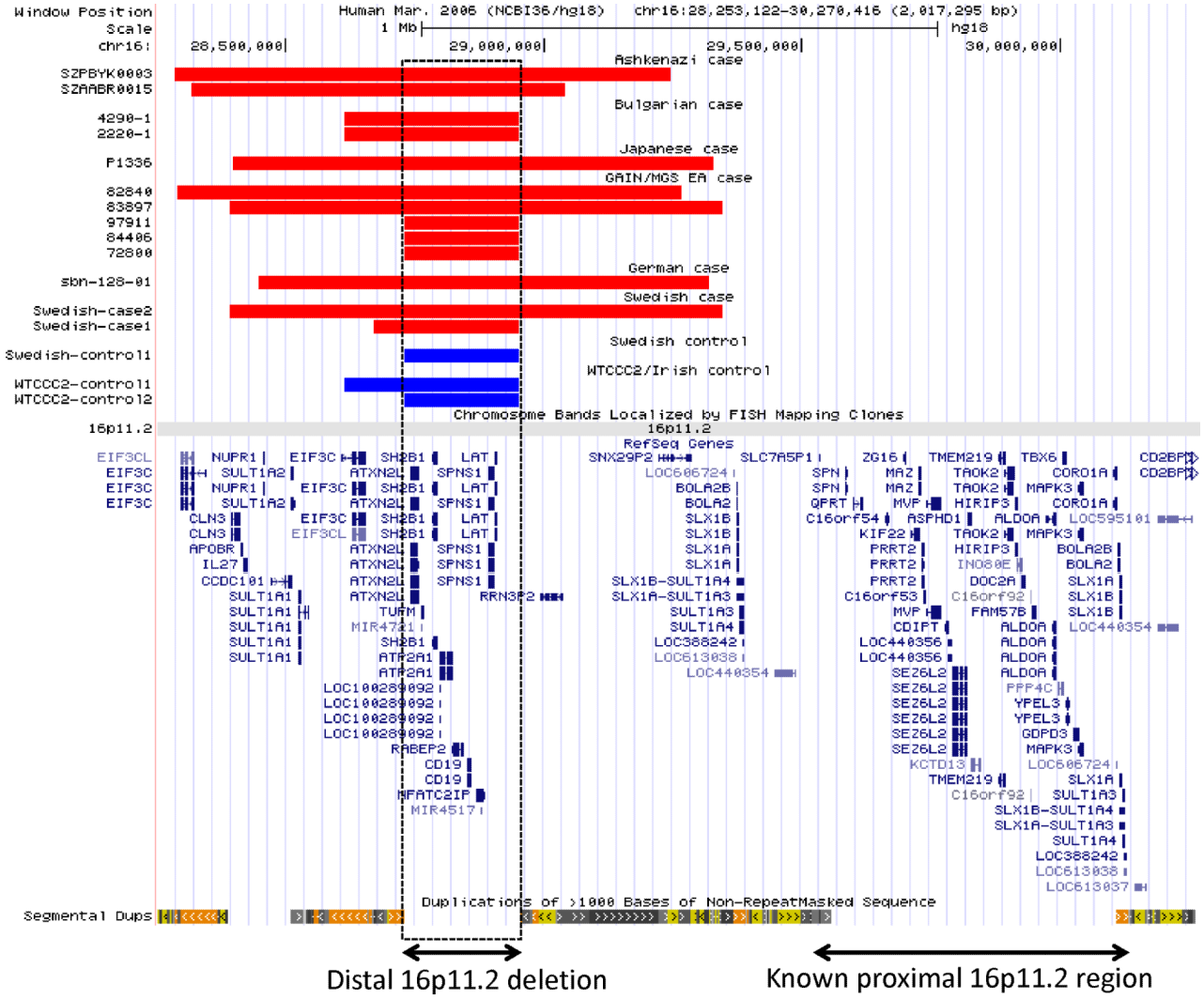

Figure 1. Microdeletions at 16p11.2 distal region in the current study

The region intersects nine genes and is flanked by two blocks of segmental duplications (SD). Red lines represent SZ cases deletions and blue lines represent control deletions. The double headed arrows indicate the intervals implicated in the current study (distal 16p11.2 deletion; minimal common region 28.73–28.95Mb, hg18), and the known 16p11.2 CNV locus (known proximal 16p11.2 region, minimal common region 29.56–30.11Mb).

Table 1

CNVs at 16p11.2 in the discovery and replication sets

Details on all samples are presented in the eSupplement, Section 1. Significance for the rate of deletions in the replication sample: Fisher one tailed test: 0.018, Cochran Mantel Haenszel test: 0.04. Significance for the total sample (Discovery + Replication): Fisher two tailed test: 0.0014, Cochran Mantel Haenszel test: 0.0031. Odds ratio (OR) for replication samples: 4.35 (95% CI = 1.18 - 16.06), OR for the total sample: 6.25 (95% CI = 1.78 - 21.93). The rate of duplications in the region is not significantly different between cases and controls. EA: European Americans, AA: African Americans. ISC: International schizophrenia consortium.

Study	Platform	Numbers of subjects tested		16p11.2 (del)		16p11.2 (dup)	
		Cases	Controls	Cases	Controls	Cases	Controls
Discovery samples							
Bulgarian Trios	Affy 6.0	662	662	2	0	1	0
Ashkenazi	Illumina Omni 1-Quad	790	1347	2	0	0	1
Replication							
GAIN/MGS EA	Affy 6.0	2671	2648	5	0	1	4
GAIN/MGS AA	Affy 6.0	1274	963	0	0	0	0
ISC	Affy 6.0/5.0	3045	3185	0	0	2	0
Vacic et al. [15]	aCGH, Nimblegen HD2	757	742	0	0	1	0
WTCCC2/Irish	Affy 6.0/ Illumina1M	1269	6175	0	2	0	4
Japanese	Affy 5.0	490	516	1	0	0	0
Swedish	Affy 6.0	1506	2089	2	1	0	2
German	Illumina (HH550, 610, 660)	1386	1627	1	0	1	2
Replication samples		**12398**	**17945**	**9 (0.073%)**	**3 (0.017%)**	**5 (0.04%)**	**12 (0.067%)**
Total (Discovery + Replication)		**13850**	**19954**	**13 (0.094%)**	**3 (0.015%)**	**6 (0.043%)**	**13 (0.065%)**

JAMA Psychiatry. Author manuscript; available in PMC 2013 August 23.